

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-078611

(43)Date of publication of application : 14.03.2000

(51)Int.CI.

H04N 13/00

(21)Application number : 10-245907

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 31.08.1998

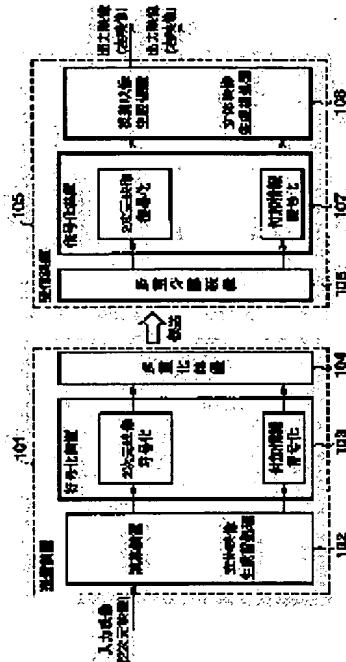
(72)Inventor : SUGIYAMA TORU  
YOSHIDA RITSUO

## (54) STEREOSCOPIC VIDEO IMAGE RECEIVER AND STEREOSCOPIC VIDEO IMAGE SYSTEM

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To generate a stereoscopic video image by decreasing the cost of the stereoscopic receiver and processing the received video information in real time.

**SOLUTION:** A stereoscopic video image transmitter 101 in the processing of generating a stereoscopic video image from a 2-dimensional video image conducts pre-processing that extracts additional information such as each pixel or a depth of each pixel in the 2-dimensional video image required to generate the stereoscopic video image from the 2-dimensional video image, and sends a coded signal consisting of the additional information obtained by the pre-processing and the 2-dimensional video image to a stereoscopic video image receiver 105. The stereoscopic video image receiver 105 receives the transmitted signal to decode respectively the 2-dimensional video image and the additional information and generates a stereoscopic video image by using parallax information based on the decoded 2-dimensional video image and the decoded additional information.



### LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C) 1998,2003 Japan Patent Office

## \* NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

CLAIMS

## [Claim(s)]

[Claim 1] In the 3-dimensional scenography receiving set which generates 3-dimensional scenography based on the image information transmitted from the sending set A receiving means to receive the signal with which the additional information for generating the two-dimensional image information and the 3-dimensional scenography which are transmitted from said sending set was encoded, A decryption means to decrypt said two-dimensional image information and said additional information from the signal received by this receiving means, respectively, The 3-dimensional scenography receiving set characterized by providing a generation means to generate the 3-dimensional scenography using parallax information, using said two-dimensional image information with which it was decrypted by this decryption means, and said additional information.

[Claim 2] Said additional information is a 3-dimensional scenography receiving set according to claim 1 characterized by being a depth value corresponding to each pixel of a two-dimensional image.

[Claim 3] The 3-dimensional scenography receiving set according to claim 2 characterized by having a means to change said depth value based on the value specified as an observer on the basis of said depth value.

[Claim 4] In the 3-dimensional scenography receiving set which generates 3-dimensional scenography based on the image information transmitted from a sending set The two-dimensional image information which is transmitted from said sending set and which was divided for every element, A receiving means to receive the signal with which the positional information which shows the physical relationship between each element, and the depth value of each element were encoded, A decryption means to decrypt said two-dimensional image information, positional information, and a depth value from the signal received with this receiving means, respectively, The 3-dimensional scenography receiving set characterized by providing a generation means to compound each element and to generate the 3-dimensional scenography using parallax information, based on said two-dimensional image information for every element decrypted with this decryption means, the positional information between each element, and the depth value for every element.

[Claim 5] The 3-dimensional scenography receiving set according to claim 4 characterized by having a means to change said depth value based on the value specified as an observer on the basis of said depth value received with said receiving means.

[Claim 6] In the solid visual system which consists of a sending set which transmits image information, and a receiving set which receives the image information transmitted from this sending set, and generates 3-dimensional scenography said sending set While encoding a pretreatment means to generate the additional information for judging the three-dimension configuration of an image from a two-dimensional video signal as pretreatment, and generating 3-dimensional scenography, and said two-dimensional video signal A coding means to encode the additional information generated by said pretreatment means, It consists of a transmitting means to transmit the signal which encoded a two-dimensional video signal and additional information with this coding means to said receiving set. Said receiving set A decryption means to decrypt said two-dimensional video signal and additional information, respectively from the signal received with a receiving means to receive the signal transmitted from said sending set, and this receiving means, The solid visual system characterized by consisting of a generation means to generate the 3-dimensional scenography using parallax information, based on said two-dimensional video signal and additional information which were decrypted with this decryption means.

[Claim 7] In the solid visual system which consists of a sending set which transmits image information, and a receiving set which receives the image information transmitted from this sending set, and generates 3-dimensional scenography said sending set While encoding the two-dimensional video signal for every element which divided the two-dimensional video signal for every element according to the contents of the image as

pretreatment, and was divided pretreatment means to judge the depth value of each element, and this pretreatment means, and the positional information between each element With a coding means to encode the depth value for every element, and this coding means, the two-dimensional video signal for said every element, It consists of a transmitting means to transmit the signal which encoded the positional information between each element, and the depth value for every element. Said receiving set A decryption means to decrypt the two-dimensional video signal for said every element, the positional information between each element, and the depth value for every element, respectively from the signal received with a receiving means to receive the signal transmitted from said sending set, and this receiving means, The solid visual system characterized by consisting of a generation means to compound each element and to generate the 3-dimensional scenography using parallax information, based on the two-dimensional video signal for said every element decrypted with this decryption means, the positional information between each element, and the depth value for every element.

---

[Translation done.]

**\* NOTICES \***

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION****[Detailed Description of the Invention]****[0001]**

[Field of the Invention] This invention relates to the 3-dimensional scenography receiving set and solid visual system which generate 3-dimensional scenography based on additional information and two-dimensional video signals, such as depth information extracted from the two-dimensional video signal.

**[0002]**

[Description of the Prior Art] Conventionally, while development of a noncommercial solid graphic display device (solid display) progresses, there are few amounts of the noncommercial software for 3-dimensional scenography, and in order to display 3-dimensional scenography, it is necessary to make 3-dimensional scenography newly.

[0003] Then, what changes the conventional two-dimensional image into 3-dimensional scenography is proposed as an approach of harnessing the property of the conventional two-dimensional image. The configuration which changes a two-dimensional image into a 3D scenography is shown in drawing 10. By the same approach as transmitting a two-dimensional video signal, a sending set 1001 encodes a two-dimensional video signal with coding equipment 1002, and is transmitted. After a receiving set 1003 receives the signal encoded from the sending set 1001 and decrypts a two-dimensional video signal with decryption equipment 1004, it generates two parallax video signals with the parallax image generation vessel 1005. That is, all the transform processing to 3-dimensional scenography is taken charge of with a receiving set 1003.

[0004] As an approach of generating 3-dimensional scenography, it assumes, for example that it is what always has near the core of a screen in this side in the depth direction, and there is the approach of generating 3-dimensional scenography so that near a core can always be seen to the front.

[0005] moreover, the image which assumes to be what has the field currently moved as other approaches in this side, and a motion gives from the left at a right eye in the case of the right -- a left eye -- receiving -- number field \*\*\*\*\* -- \*\* -- it is carrying out and the field currently moved will be perceived to the front.

[0006] Research which current and the above image processing techniques accomplish development, and are not concerned with an animation and a still picture, but presumes a three-dimension configuration from a two-dimensional image is done briskly. On the other hand, standardization of MPEG4 is made about coding of an image. The configuration is shown in drawing 11. First, a two-dimensional image as shown in drawing 12 (a) is inputted into a sending set 1011. In an arithmetic unit 1012, as shown in drawing 12 (b) - (d), while disassembling an image into a component and encoding each element with coding equipment 1013, the information on the physical relationship between each element is also encoded. Multiplex [ of each encoded signal ] is carried out with multiplexer 1014, and it is transmitted from a sending set (transmission). It is received by the receiving set 1015 and this transmitted signal is first separated for every element by the demultiplexer 1016. Each separated element is decrypted by decryption equipment 1017, and also decrypts the information on the physical relationship between elements to coincidence. As shown in drawing 12 (e), the decrypted signal is compounded based on the information on physical relationship in the synthetic section 1018, and is outputted as a subject-copy image and same image.

**[0007]**

[Problem(s) to be Solved by the Invention] However, the output of the equipment of above-mentioned drawing 11 is a two-dimensional image to the last. Then, when a 3-dimensional scenography receiving set tends to perform altogether processing which changes a two-dimensional image into the 3D scenography which has real nature more, there is a fault that a 3-dimensional scenography receiving set serves as a large sum, or it

becomes and difficult to process throughput on real time.

[0008] Then, this invention can lower the cost of a 3-dimensional scenography receiving set, and aims at offering the 3-dimensional scenography receiving set and solid visual system which can process the received image information on real time, and can generate 3-dimensional scenography.

[0009]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, the 3-dimensional scenography receiving set of this invention In what generates 3-dimensional scenography based on the image information transmitted from the sending set A receiving means to receive the signal with which the additional information for generating the two-dimensional image information and the 3-dimensional scenography which are transmitted from said sending set was encoded, It constitutes from a generation means to generate the 3-dimensional scenography using parallax information, using a decryption means to decrypt said two-dimensional image information and additional information from the signal received by this receiving means, respectively, and said two-dimensional image information, with which it was decrypted by this decryption means and said additional information.

[0010] Since the additional information which did not need to make anew the information for making 3-dimensional scenography using two-dimensional image information from the receiving set, and has been sent by the above-mentioned means since two-dimensional image information and additional information are sent from a sending set can be used immediately, signal processing in a receiving set can be simplified greatly.

[0011]

[Embodiment of the Invention] Hereafter, the gestalt of implementation of this invention is explained with reference to a drawing. Drawing 1 is a block diagram for explaining the gestalt of implementation of the 1st of this invention. This solid visual system consists of a sending set (3-dimensional scenography sending set) which transmits image information, and a receiving set (3-dimensional scenography receiving set) which receives the image information transmitted from this sending set, and displays that image. For example, as a sending set, they are the sending set in the broadcasting station of television, a video camera, or a video tape recorder (VTR). As a receiving set which receives and displays the image information transmitted from such a sending set, there are three dimensional display equipments, such as a solid display.

[0012] As shown in drawing 1, an arithmetic unit 102, coding equipment 103, and multiplexer 104 constitute the sending set 101 of a transmitting side, and it constitutes the receiving set 105 of a receiving side with a demultiplexer 106, decryption equipment 107, and parallax image generation equipment 108.

[0013] First, the two-dimensional video signal which should be transmitted to the arithmetic unit 102 in the sending set 101 of a transmitting side is inputted. In an arithmetic unit 102, pretreatment required for generation of 3-dimensional scenography is performed, and additional information effective in generation of 3-dimensional scenography is computed. As useful information, the three-dimension configuration directly presumed from the two-dimensional image is in generation of 3-dimensional scenography, and this shows with the depth value over each pixel of for example, a two-dimensional image.

[0014] There is how a three-dimension configuration is extracted [ much ] from a two-dimensional image in the field of current and a computer vision, and research is done. Since the two-dimensional video signal which changes in the direction of a time-axis is inputted into the above-mentioned arithmetic unit 102 of the gestalt of this 1st operation, as a parameter when presuming a three-dimension configuration, it is arrangement, magnitude, etc. of motion parallax, shading, the inclination of a texture, line distance, the two-dimensional configuration of an element, the lap between elements, and an element. Furthermore, there is a variation rate between motion vectors etc. as a parameter when presuming motion parallax, the two-dimensional configuration of an element, and the lap between elements, and the above-mentioned parameter relates each other to each other. Therefore, all the above-mentioned parameters will correspond besides a depth value as additional information effective in 3-dimensional scenography generation. That is, the parameter of a three-dimension element indispensable in order to generate 3-dimensional scenography judges with a sending set 101, and transmits to a receiving set 105.

[0015] Although the operation of the additional information involved in presumption of the three-dimension configuration from the above two-dimensional images needs many processings, with the gestalt of this 1st operation, in order to give these processings to a transmitting side, real time nature is not needed in many cases, and the high arithmetic unit of capacity is used at a transmitting side compared with a receiving side. For this reason, generation of 3-dimensional scenography with a more high precision is attained. Therefore, although it is desirable to presumption of the three-dimension configuration from a two-dimensional image for the

operation which gave [ above-mentioned.] explanation to generate automatically, people double with a two-dimensional image manually, depth is set to it, and it is good for it also considering it as additional information. [0016] Moreover, the additional information generated with the above-mentioned arithmetic unit 102 is encoded with coding equipment 103 with two-dimensional image information. These two encoded signals are multiplexed with multiplexer 104, and are transmitted to a receiving set 105 from a sending set 101.

[0017] And a receiving set 105 receives the signal transmitted from the sending set 101. First, it separates into a two-dimensional video signal and an additional information signal by the demultiplexer 106 of a receiving set 105, and this signal decrypts each signal with decryption equipment 107. The decrypted signal is offered to an observer's separate eye, after inputting into parallax image generation equipment 108, performing after treatment required for generation of 3-dimensional scenography and generating two parallax images.

[0018] Drawing 2 shows the conceptual diagram at the time of using the depth value for every pixel as additional information. Drawing 2 (a) As shown in - (d), the two-dimensional video signal 111 inputted into a sending set 101 and the depth value (additional information) 112 for every pixel computed from this two-dimensional video signal 111 are transmitted to a receiving set 105. In a receiving set 105, the parallax image (left image) 113 and the parallax image (right image) 114 which were generated from the received information are outputted outside.

[0019] Moreover, when the depth value for every pixel is known, it can be uniquely found from the magnitude of a screen, an observer's observation distance, and the distance between both eyes into which location of a parallax image each pixel of a two-dimensional image is changed.

[0020] For example, as shown in drawing 3 , the pixel of the original location of a two-dimensional image is moved to the pixel location for right eyes, and the pixel location for left eyes according to the depth value, and geometry is asked for a location. Moreover, in case a two-dimensional image is changed into a parallax image, it is necessary to generate the pixel which does not exist in a two-dimensional image. In this case, for example, it interpolates and generates from a contiguity pixel.

[0021] As stated above, in case 3-dimensional scenography is generated from a two-dimensional image, a transmitting side and a receiving side share processing. The high depth value of a processing load is judged by the transmitting side on non-real time, and a parallax image is generated by the receiving side from the depth value which needs to be changed with the configuration and observation distance of a display.

[0022] Thereby, generation of 3-dimensional scenography with a high precision in a receiving set is attained on real time. Moreover, since a depth value is sent as additional information of the usual two-dimensional video signal, television which can display only a two-dimensional image can maintain compatibility.

[0023] Next, drawing 4 is a block diagram for explaining the gestalt of implementation of the 2nd of this invention. First, the two-dimensional video signal which should be transmitted to the arithmetic unit 202 in the sending set 201 of a transmitting side is inputted. In an arithmetic unit 202, pattern recognition of the inputted two-dimensional video signal is carried out, and field division of the screen of an image is carried out for every element. The technique of field division is for example, an image processing handbook (Showa 62 issue \*\*\*\* Co.) etc., and various technique is introduced. Moreover, in an arithmetic unit 202, processing which presumes the depth for every element is performed to coincidence.

[0024] The video signal decomposed for every element with the arithmetic unit 202 is encoded for every element with coding equipment 203. In addition, although this coding equipment 203 is shown that two or more coding processing sections are arranged inside, and it processes to juxtaposition as shown in drawing 4 , it may process two or more elements in the one coding processing section sequentially.

[0025] Coding equipment 203 also encodes the physical relationship information by what kind of arrangement each element is compounded while encoding each element. The scene description language (BIFS:Binary Format for Scenes) standardized by MPEG4 may be used for the technique of this coding.

[0026] Moreover, coding equipment 203 also encodes the depth value for every element as additional information. An example of this coding format is shown in drawing 5 . As shown in drawing 5 , an index (index) also stores the value which shows the die length of a signal by the identification code which shows that the signal is the information on a depth value. An identifier (name) is the identification code for every element. The depth (depth) shows the depth value of the element. Here, although the information on a depth value is treated as additional information other than positional information, since arrangement of the three-dimension space of each element can be specified also by the above-mentioned scene description language, a depth value may be specified within a scene description language.

[0027] And after multiplexing each signal encoded with coding equipment 203 with multiplexer 204 and changing

it into a transmission format, it is transmitted from transmission equipment 101 (transmission).

[0028] The demultiplexer 206 in a receiving set 205 receives, and this transmitted signal is separated for each [which is multiplexed] element of every. Each of this separated element is decrypted in decryption equipment 207, and also decrypts the additional information which shows physical relationship and the depth value of each element to coincidence.

[0029] As shown in drawing 4, decryption equipment 207 arranges two or more decryption processing sections inside, processes a decryption to juxtaposition, but as long as it can process this processing within real time, it may process two or more elements in the one decryption processing section sequentially.

[0030] After being inputted into the parallax image generation machine 208, and parallax image generation equipment's 208 performing after treatment required for 3-dimensional scenography generation and generating two parallax images, an observer's separate eye is provided with the signal decrypted by decryption equipment 207.

[0031] Drawing 6 is a conceptual diagram at the time of using the depth value for every element as additional information. That is, to the subject-copy image 221 of a two-dimensional image, as shown in drawing 6 (a), as shown in drawing 6 R> 6 (b) – (d), it divides into the image 222 of the background in the subject-copy image 221, the image 223 of a house, and the image 224 of a bus. As depth information on each of this divided element and each element, a depth value is determined from the back in order of the image 222 of a background, the image 223 of a house, and the image 224 of a bus. And a subject-copy image encodes the physical relationship information on each of these elements, and the depth value of each element, and it transmits to a receiving set. A receiving set decrypts the two-dimensional video signal encoded and transmitted, the physical relationship information between each element, and the depth value of each element, and based on those information, as shown in drawing 6 (e) and (f), it generates the right image 225 as a parallax image, and the left image 226.

[0032] Next, in case a parallax image is changed, the interpolation approach of a field without the pixel to produce is shown in drawing 7. That is, in order to take out a steric effect, when a photographic subject image is rotated or moved and a stereoscopic model is created, the original two-dimensional screen product and the area for a stereoscopic model are not necessarily in agreement. For this reason, the non-picture area which does not have a pixel to a background image may arise.

[0033] For example, as shown in drawing 7 (a) and (b), a receiving set receives each element disassembled into the image 232 of a background, and the image 233 of a house to the subject-copy image 231 with the sending set, and a parallax image is generated according to the image and depth value of each of that element. Then, as shown in drawing 7 (c), a non-picture area without a pixel may occur. In this case, as shown in drawing 7 (d), it interpolates by expanding the image 233 of a house by space shaft orientations. Moreover, it is also possible to interpolate a field without a pixel from a contiguity pixel.

[0034] Drawing 8 is a block diagram for explaining the gestalt of implementation of the 3rd of this invention. In addition, the same part as drawing 1 attaches the same sign, and omits the explanation. With the gestalt of this operation, as shown in drawing 8, in a receiving set 301, parallax image generation equipment 302 generates 3-dimensional scenography based on the parallax information which the user inputted with the parallax adjustment input unit 303 which an observer adjusts. It enables this to offer the 3-dimensional scenography doubled with liking of an observer. That is, a user can adjust parallax information and can adjust the feeling of depth of a stereoscopic model.

[0035] Drawing 9 (a) – (c) is drawing for explaining an example of parallax adjustment. As shown in drawing 9 (a), the depth value 312 is specified as a receiving set 301 by pretreatment of the arithmetic unit 102 of a sending set 101 to the screen side 311.

[0036] And as shown in drawing 9 (b), when a user adjusts the datum level of depth, a user inputs the adjustment value of the datum level of depth with the parallax adjustment input unit 303. Then, the parallax image generation equipment 302 of a receiving set 301 changes the datum level of a depth value into the depth value 313 from the depth value 312, and generates a parallax image. The depth value specified by the sending set 101 can be changed by this according to the adjustment value of the datum level of the depth which an observer inputs, the image generated can be changed into the image which is visible to the image which is visible to the whole at back, or this side, and the 3-dimensional scenography according to liking of an observer can be offered.

[0037] Moreover, when a user adjusts the amount of depth, a user inputs the amount of adjustments of the amount of depth with the parallax adjustment input unit 303. Then, the parallax image generation equipment 302

of a receiving set 301 makes the depth value 312 fluctuate according to the amount of adjustments, is changed into the depth value 314, and generates a parallax image. The image which changes and generates by this the depth value specified by the sending set according to the amount of adjustments of the amount of depth which an observer inputs can be changed into the large image of depth perception, or the small image of depth perception, and the 3-dimensional scenography according to liking of an observer can be offered.

[0038] above-mentioned the 1- as the gestalt of each 3rd operation explained, a 3-dimensional-scenography sending set performs pretreatment which extracts additional information, such as each pixel in a required two-dimensional image, or a depth value of each element, when generating 3-dimensional scenography from a two-dimensional image, and transmits it to a 3-dimensional-scenography receiving set in the processing which generates 3-dimensional scenography from a two-dimensional image as a signal which encoded the additional information acquired with this pretreatment, and a two-dimensional image. A 3-dimensional scenography receiving set receives the transmitted signal, decrypts a two-dimensional image and additional information, respectively, and generates the 3-dimensional scenography using parallax information based on the two-dimensional image and additional information which were decrypted.

[0039] A sending set can perform processing which extracts by this additional information, such as presumption of a depth value which requires many operations, on non-real time, additional information with a high precision can be acquired, and with a receiving set, when image information is received, highly precise 3-dimensional scenography can be generated on real time.

[0040] Moreover, since an observer can change additional information, such as a depth value, by the receiving side in case 3-dimensional scenography generates based on additional information, such as a depth value acquired by the transmitting side, offer of the 3-dimensional scenography doubled with liking of an observer is attained.

[0041]

[Effect of the Invention] As explained in full detail above, according to this invention, the cost of a 3-dimensional scenography receiving set can be lowered, the received image information can be processed on real time, and 3-dimensional scenography can be generated.

---

[Translation done.]

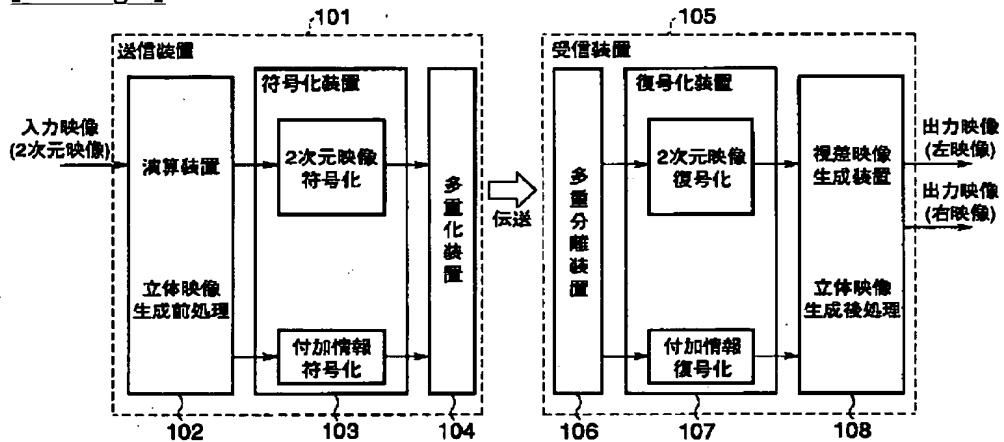
## \* NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

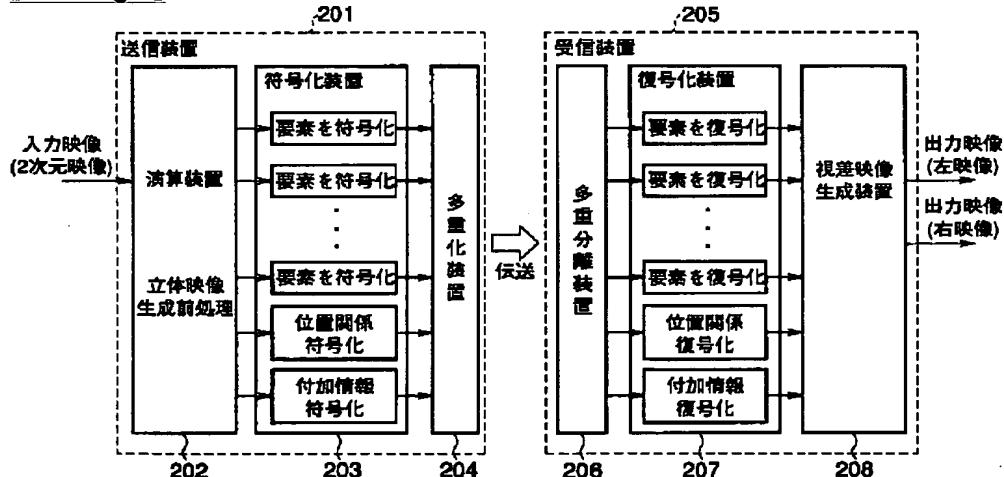
1. This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

## DRAWINGS

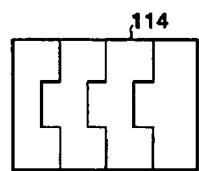
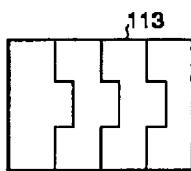
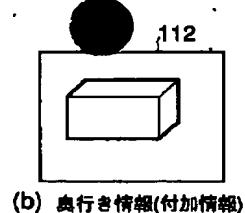
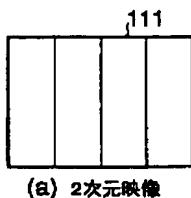
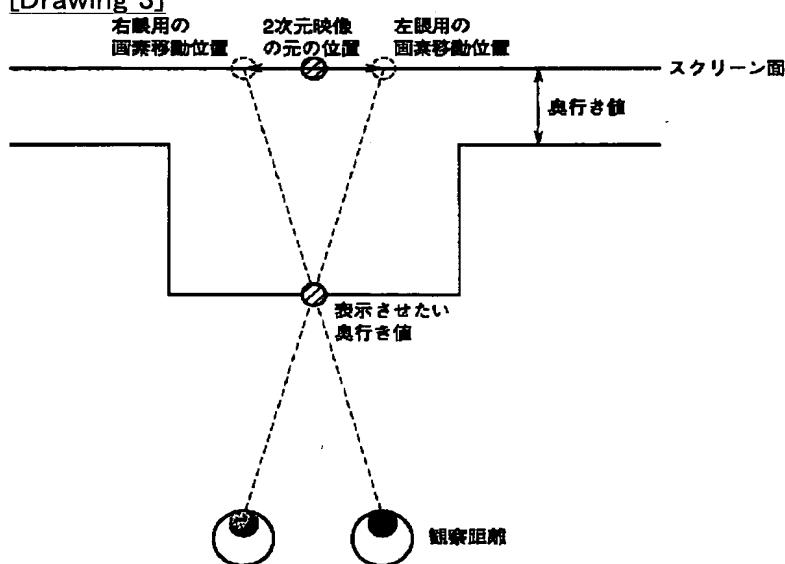
## [Drawing 1]



## [Drawing 4]



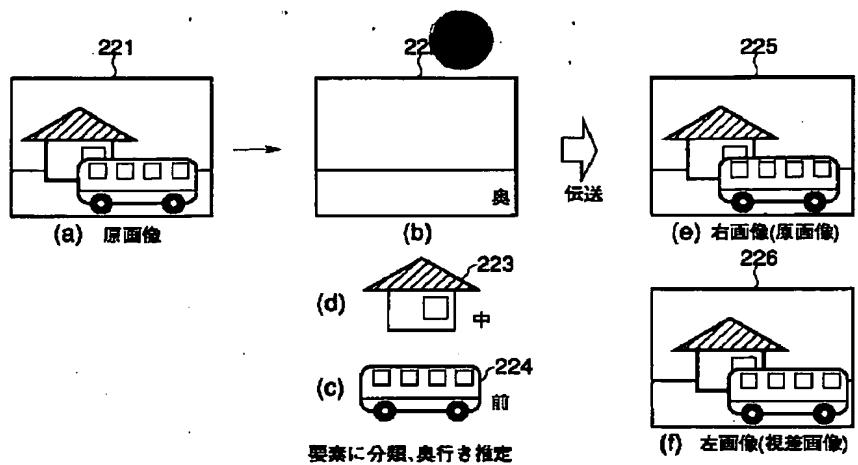
## [Drawing 2]

[Drawing 3][Drawing 5]

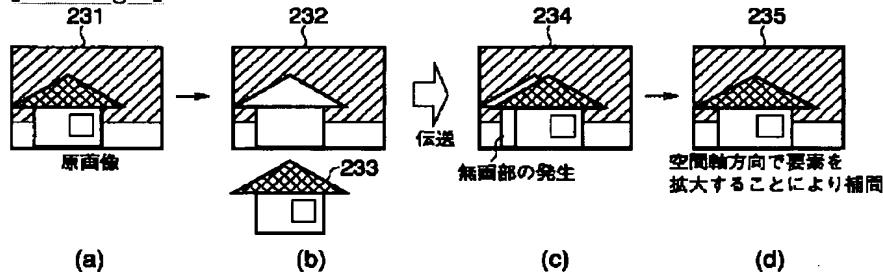
211

Index	name	depth	name	depth	.....	name	depth
-------	------	-------	------	-------	-------	------	-------

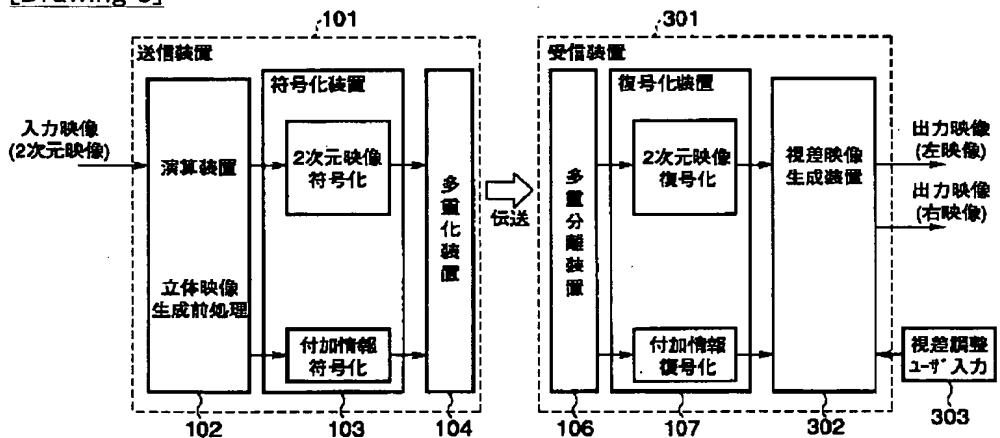
[Drawing 6]



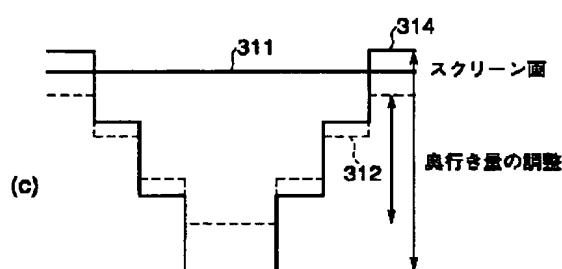
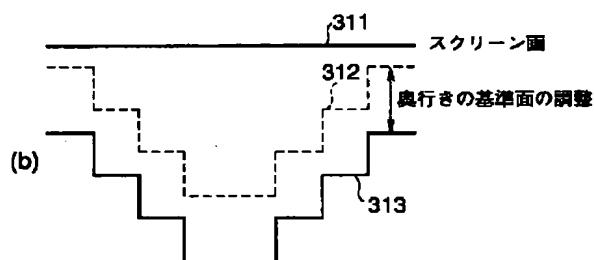
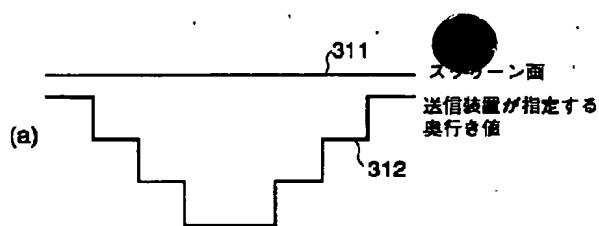
[Drawing 7]



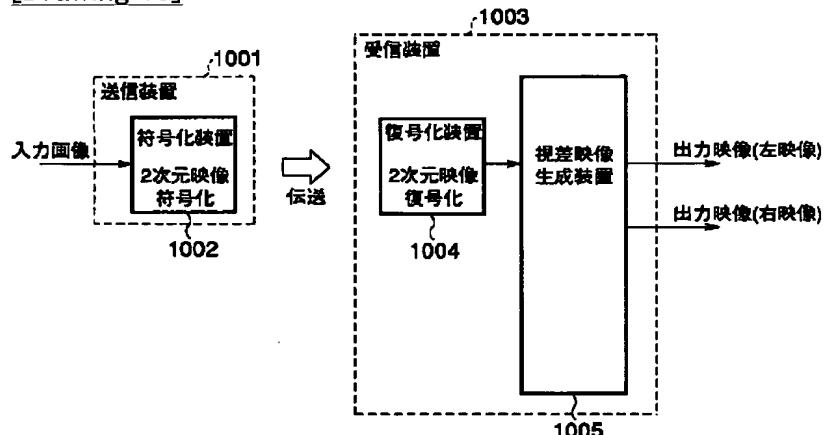
[Drawing 8]



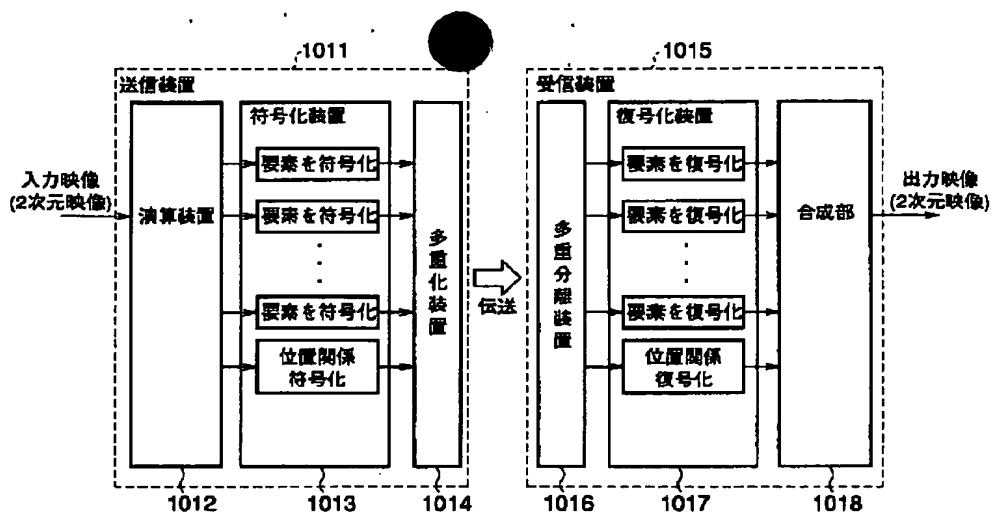
[Drawing 9]



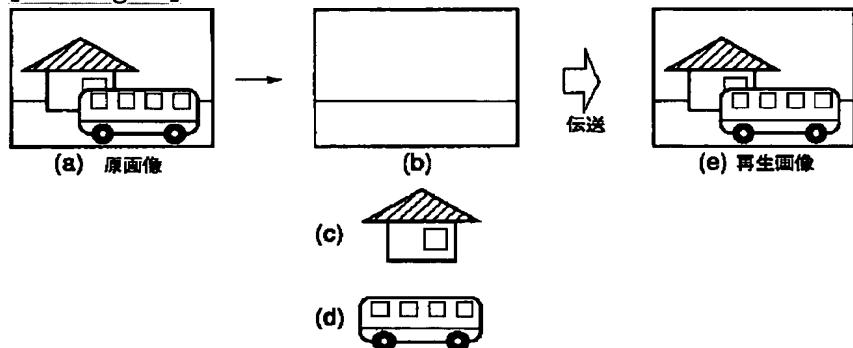
[Drawing 10]



[Drawing 11]



[Drawing 12]



[Translation done.]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-78611

(P2000-78611A)

(43)公開日 平成12年3月14日 (2000.3.14)

(51) Int.Cl.<sup>7</sup>  
H 0 4 N 13/00

識別記号

F I  
H 0 4 N 13/00

テマコード<sup>\*</sup> (参考)  
5 C 0 6 1

審査請求 未請求 請求項の数 7 O L (全 10 頁)

(21)出願番号 特願平10-245907

(22)出願日 平成10年8月31日 (1998.8.31)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 杉山 徹

神奈川県横浜市磯子区新杉田町8番地 株

式会社東芝マルチメディア技術研究所内

(72)発明者 吉田 律生

神奈川県横浜市磯子区新杉田町8番地 株

式会社東芝マルチメディア技術研究所内

(74)代理人 100058479

弁理士 鈴江 武彦 (外6名)

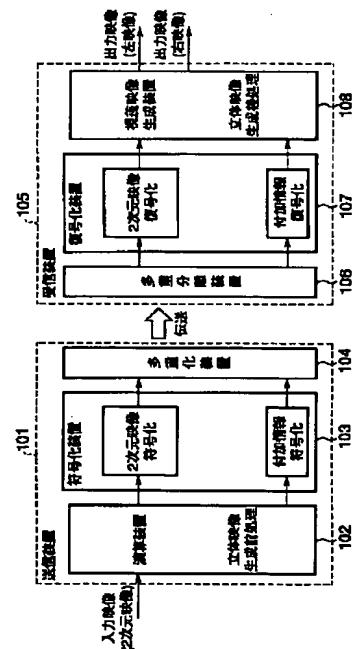
F ターム(参考) 5C061 AA29 AB01 AB08 AB11 AB12  
AB24

(54)【発明の名称】 立体映像受信装置及び立体映像システム

(57)【要約】

【課題】 立体映像受信装置のコストを下げるとともに、受信された映像情報をリアルタイムで処理して立体映像を生成する。

【解決手段】 2次元映像から立体映像を生成する処理において、立体映像送信装置101は、2次元映像から立体映像を生成する上で必要な2次元映像における各画素あるいは各要素の奥行き値等の付加情報を抽出する前処理を行い、この前処理で得られた付加情報と2次元映像とを符号化した信号として立体映像受信装置105へ送信する。立体映像受信装置105は、送信された信号を受信して2次元映像と付加情報とをそれぞれ復号化し、復号化された2次元映像と付加情報とに基づいて、視差情報を用いた立体映像を生成する。



## 【特許請求の範囲】

【請求項1】送信装置から送信された映像情報に基づいて、立体映像を生成する立体映像受信装置において、前記送信装置から送信される2次元映像情報と立体映像を生成するための付加情報とが符号化された信号とを受信する受信手段と、

この受信手段により受信された信号から前記2次元映像情報と前記付加情報をそれぞれ復号化する復号化手段と、

この復号化手段により復号化された前記2次元映像情報と前記付加情報を用いて、視差情報を用いた立体映像を生成する生成手段と、を具備したことを特徴とする立体映像受信装置。

【請求項2】前記付加情報は、2次元映像の各画素に対応する奥行き値であることを特徴とする請求項1記載の立体映像受信装置。

【請求項3】前記奥行き値を基準として、観察者に指定される値に基づいて前記奥行き値を変更する手段を有することを特徴とする請求項2記載の立体映像受信装置。

【請求項4】送信装置から送信される映像情報に基づいて、立体映像を生成する立体映像受信装置において、前記送信装置から送信される要素毎に分割された2次元映像情報と、各要素間の位置関係を示す位置情報と、各要素の奥行き値とが符号化された信号とを受信する受信手段と、

この受信手段により受信した信号から前記2次元映像情報、位置情報、及び奥行き値をそれぞれ復号化する復号化手段と、

この復号化手段により復号化した各要素毎の前記2次元映像情報、各要素間の位置情報、及び各要素毎の奥行き値に基づいて、各要素を合成して視差情報を用いた立体映像を生成する生成手段と、を具備したことを特徴とする立体映像受信装置。

【請求項5】前記受信手段により受信した前記奥行き値を基準として、観察者に指定される値に基づいて前記奥行き値を変更する手段を有したこととを特徴とする請求項4記載の立体映像受信装置。

【請求項6】映像情報を送信する送信装置と、この送信装置から送信される映像情報を受信して立体映像を生成する受信装置とからなる立体映像システムにおいて、前記送信装置は、

前処理として2次元映像信号から像の3次元形状を判断して立体映像を生成するための付加情報を生成する前処理手段と、

前記2次元映像信号を符号化するとともに、前記前処理手段により生成された付加情報を符号化する符号化手段と、

この符号化手段により2次元映像信号と付加情報を符号化した信号を前記受信装置に送信する送信手段とから

なり、

前記受信装置は、

前記送信装置から送信される信号を受信する受信手段と、

この受信手段により受信した信号から前記2次元映像信号と付加情報をそれぞれ復号化する復号化手段と、この復号化手段により復号化した前記2次元映像信号と付加情報を基づいて、視差情報を用いた立体映像を生成する生成手段とからなることを特徴とする立体映像システム。

【請求項7】映像情報を送信する送信装置と、この送信装置から送信される映像情報を受信して立体映像を生成する受信装置とからなる立体映像システムにおいて、前記送信装置は、

前処理として2次元映像信号を映像の内容に従って要素毎に分割して、各要素の奥行き値を判断する前処理手段と、

この前処理手段により分割された各要素毎の2次元映像信号及び各要素間の位置情報を符号化するとともに、各要素毎の奥行き値を符号化する符号化手段と、

この符号化手段により前記各要素毎の2次元映像信号、各要素間の位置情報、及び各要素毎の奥行き値を符号化した信号を送信する送信手段とからなり、

前記受信装置は、

前記送信装置から送信された信号を受信する受信手段と、

この受信手段により受信した信号から前記各要素毎の2次元映像信号、各要素間の位置情報、及び各要素毎の奥行き値をそれぞれ復号化する復号化手段と、

この復号化手段により復号化した前記各要素毎の2次元映像信号、各要素間の位置情報および各要素毎の奥行き値に基づいて、各要素を合成して視差情報を用いた立体映像を生成する生成手段とからなることを特徴とする立体映像システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、2次元映像信号から抽出された奥行き情報等の付加情報を用いて立体映像を生成する立体映像受信装置及び立体映像システムに関する。

【0002】

【従来の技術】従来、民生用の立体映像表示装置（立体ディスプレイ）の開発が進む一方で民生用の立体映像用ソフトの量は少なく、立体映像を表示するためには新規に立体映像を制作する必要がある。

【0003】そこで、従来の2次元映像の資産を活かす方法として、従来の2次元映像を立体映像に変換するものが提案されている。2次元映像を3次元映像に変換する構成を図10に示す。送信装置1001は、2次元映像信号を伝送するのと同じ方法で2次元映像信号を符号

化装置1002で符号化し送信する。受信装置1003は、送信装置1001からの符号化された信号を受信して復号化装置1004で2次元映像信号を復号化した後、視差映像生成器1005で2つの視差映像信号を生成する。つまり、受信装置1003で立体映像への変換処理の全てを受け持つ。

【0004】立体映像を生成する方法としては、例えば、画面の中心付近が奥行き方向で常に手前にあるものと仮定し、中心付近が常に手前に見えるように立体映像を生成する方法がある。

【0005】また、他の方法としては動いている領域が手前にあるものと仮定し、動きが左から右方向の場合は、右眼に与える映像は左眼に対して数フィールド遅れたものとすることで、動いている領域が手前に知覚されることになる。

【0006】現在、上記のような映像処理技術が発展を遂げ、動画、静止画に関わらず2次元映像から3次元形状を推定する研究が盛んに行われている。一方、映像の符号化に関してはMPEG4の規格化がなされている。その構成を図11に示す。まず、図12(a)に示すような2次元映像が送信装置1011に入力される。演算装置1012では、図12(b)～(d)に示すように、画像を構成要素に分解し、符号化装置1013で各要素の符号化を行なうとともに各要素間の位置関係の情報も符号化する。符号化した各信号は多重化装置1014で多重され送信装置より送信(伝送)される。この送信された信号は受信装置1015により受信され、まず、多重分離装置1016で各要素毎に分離される。分離した各要素は復号化装置1017により復号化され、同時に、要素間の位置関係の情報も復号化する。復号化した信号は、図12(e)に示すように、合成部1018にて位置関係の情報に基づいて合成され、原画像と同様な映像として出力される。

【0007】

【発明が解決しようとする課題】しかし、上記図11の装置は、あくまでも出力が2次元映像である。そこで、2次元映像をよりリアル性のある3次元映像に変換する処理を立体映像受信装置で全て行おうとすると、立体映像受信装置が高額となったり、処理量が多くリアルタイムでの処理が困難となるという欠点がある。

【0008】そこで、本発明は、立体映像受信装置のコストを下げることができ、受信された映像情報をリアルタイムで処理して立体映像を生成することができる立体映像受信装置及び立体映像システムを提供することを目的とする。

【0009】

【課題を解決するための手段】上記した課題を解決するために、この発明の立体映像受信装置は、送信装置から送信された映像情報に基づいて、立体映像を生成するものにおいて、前記送信装置から送信される2次元映像情

報と立体映像を生成するための付加情報とが符号化された信号とを受信する受信手段、この受信手段により受信された信号から前記2次元映像情報と付加情報とをそれぞれ復号化する復号化手段、およびこの復号化手段により復号化された前記2次元映像情報と前記付加情報とを用いて、視差情報を用いた立体映像を生成する生成手段から構成する。

【0010】上記の手段により、送信装置から2次元映像情報と付加情報とが送られてくるため、受信装置では、2次元映像情報を用いて立体映像を作るための情報を改めて作る必要はなく、送られてきた付加情報をすぐに用いることができるので、受信装置での信号処理が大きく簡略化できることになる。

【0011】

【発明の実施の形態】以下、この発明の実施の形態について図面を参照して説明する。図1は、この発明の第1の実施の形態について説明するための構成図である。この立体映像システムは、映像情報を送信する送信装置(立体映像送信装置)と、この送信装置から送信される映像情報を受信してその映像を表示する受信装置(立体映像受信装置)からなる。例えば、送信装置としては、テレビジョンの放送局における送信装置、ビデオカメラ、あるいはビデオテープレコーダ(VTR)等である。このような送信装置から送信される映像情報を受信して表示する受信装置としては、立体ディスプレイ等の立体表示装置がある。

【0012】図1に示すように、送信側の送信装置101は、演算装置102、符号化装置103、多重化装置104により構成し、受信側の受信装置105は、多重分離装置106、復号化装置107、および視差映像生成装置108により構成する。

【0013】まず、送信側の送信装置101内の演算装置102に送信すべき2次元映像信号を入力する。演算装置102では立体映像の生成に必要な前処理を行い、立体映像の生成に有効な付加情報を算出する。立体映像の生成に有用な情報としては、直接的には2次元映像から推定した3次元形状があり、これは、例えば、2次元映像の各画素に対する奥行き値で示す。

【0014】2次元映像から3次元形状をいかに抽出するかは、現在、コンピュータビジョンの分野で多く研究が行われている。この第1の実施の形態の上記演算装置102には、時間軸方向で変化する2次元映像信号が入力されるため、3次元形状を推定する上のパラメータとしては運動視差、陰影、テクスチャの勾配、線遠近、要素の2次元形状、要素間の重なり、要素の配置および大きさ等である。さらに、運動視差、要素の2次元形状、要素間の重なりを推定する上のパラメータとして動きベクトルや動きベクトル間の変位などがあり、上記パラメータがお互いに関連し合っている。そのため、立体映像生成に有効な付加情報として、奥行き値の他にも上

記の全てのパラメータが相当することとなる。つまり、立体映像を生成するために最低限必要な3次元要素のパラメータが送信装置101で判断し、受信装置105に送信する。

【0015】上記のような2次元映像からの3次元形状の推定に絡む付加情報の演算は多くの処理を必要とするが、この第1の実施の形態では、これらの処理を送信側に持たせるためにリアルタイム性が必要とされないケースが多く、かつ送信側では受信側に比べ能力の高い演算装置を用いる。このため、より精度の高い立体映像の生成が可能になる。よって、2次元映像からの3次元形状の推定には、上記説明した演算によって自動的に生成するのが望ましいが、この他に、人が手動で2次元映像に合わせ奥行きを設定し、それを付加情報としても良い。

【0016】また、上記演算装置102にて生成した付加情報は、2次元映像情報とともに符号化装置103にて符号化する。この符号化した2つの信号は、多重化装置104にて多重化し、送信装置101から受信装置105に送信する。

【0017】そして、送信装置101から伝送された信号は、受信装置105にて受信する。この信号は、まず、受信装置105の多重分離装置106にて2次元映像信号と付加情報信号に分離し、復号化装置107にてそれぞれの信号の復号化を行う。復号化された信号は視差映像生成装置108に入力して、立体映像の生成に必要な後処理を行って、2つの視差映像が生成された後、観察者の別々の眼に対して提供する。

【0018】図2は、付加情報として各画素毎の奥行き値を用いた場合の概念図を示す。図2(a)～(d)に示すように、送信装置101に入力される2次元映像信号111と、この2次元映像信号111から算出された画素毎の奥行き値(付加情報)112とを受信装置105に伝送する。受信装置105では受信された情報から生成した視差映像(左映像)113、視差映像(右映像)114を外部に出力する。

【0019】また、各画素毎の奥行き値が既知の場合、2次元映像の各画素が視差映像との位置に変換されるかはスクリーンの大きさ、観察者の観察距離および両眼間の距離から一意的に求まる。

【0020】例えば、図3に示すように、2次元映像の元の位置の画素をその奥行き値に応じて、右眼用の画素位置および左眼用の画素位置に移動させて幾何学的配置を位置を求める。また、2次元映像を視差映像に変換する際、2次元映像に存在しない画素を生成する必要もある。この場合は、例えば、隣接画素から補間して生成する。

【0021】以上述べてきたように、2次元映像から立体映像を生成する際、送信側と受信側で処理を分担する。処理負荷の高い奥行き値の判断を非リアルタイムで送信側で行い、表示装置の形状や観察距離によって変化

させる必要がある奥行き値から視差映像の生成を受信側で行う。

【0022】これにより、受信装置での精度の高い立体映像の生成がリアルタイムで可能になる。また、奥行き値は通常の2次元映像信号の付加情報として送られてくるので、2次元映像しか表示できないテレビとも互換性が保てる。

【0023】次に、図4はこの発明の第2の実施の形態について説明するための構成図である。まず、送信側の送信装置201内の演算装置202に送信すべき2次元映像信号を入力する。演算装置202では、入力された2次元映像信号をパターン認識し、映像の画面を要素毎に領域分割する。領域分割の手法は、例えば、映像処理ハンドブック(昭和62年発行 昭晃社)などで、様々な手法が紹介されている。また、演算装置202では、同時に、各要素毎の奥行きを推定する処理を行う。

【0024】演算装置202にて要素毎に分解された映像信号は、符号化装置203にて要素毎に符号化される。なお、この符号化装置203は、図4に示すように、内部に複数個の符号化処理部を配置し並列に処理するように示されているが、1つの符号化処理部で複数の要素を順次処理しても良い。

【0025】符号化装置203は、各要素を符号化するとともに各要素をどのような配置で合成するかという位置関係情報もまた符号化する。この符号化の手法は、MPEG4で規格化されているシーン記述言語(BIFS: Binary Format for Scenes)を用いてもよい。

【0026】また、符号化装置203は、付加情報として各要素毎の奥行き値も符号化する。この符号化フォーマットの一例を図5に示す。図5に示すように、インデックス(index)は、その信号が奥行き値の情報であることを示す識別コードで信号の長さを示す値も格納する。名前(name)は、各要素毎の識別コードである。深さ(depth)は、その要素の奥行き値を示す。ここでは、奥行き値の情報が位置情報とは別の付加情報として扱われているが、上記シーン記述言語でも各要素の3次元空間の配置が指定できるのでシーン記述言語内で奥行き値を指定しても良い。

【0027】そして、符号化装置203にて符号化された各信号は、多重化装置204で多重化し、伝送フォーマットに変換された後、伝送装置201より伝送(送信)する。

【0028】この伝送された信号は、受信装置205内の多重分離装置206により受信し、多重化されている各要素毎に分離する。この分離された各要素は、復号化装置207において復号化し、同時に、位置関係および各要素の奥行き値を示す付加情報も復号化する。

【0029】復号化装置207は、図4に示すように、内部に複数個の復号化処理部を配置し、並列に復号化の

処理を行うが、この処理がリアルタイム内で処理できれば1つの復号化処理部で複数の要素を順次処理しても良い。

【0030】復号化装置207により復号化された信号は、視差映像生成器208に入力され、視差映像生成装置208により立体映像生成に必要な後処理を行い、2つの視差映像が生成された後、観察者の別々の眼に提供する。

【0031】図6は、付加情報として各要素毎の奥行き値を用いた場合の概念図である。すなわち、図6(a)に示すように、2次元映像の原画像221に対して、図6(b)～(d)に示すように、原画像221中の背景の映像222、家の映像223、バスの映像224に分割する。この分割された各要素および各要素の奥行き情報として、奥から背景の映像222、家の映像223、バスの映像224の順に奥行き値を決定する。そして、これらの各要素の位置関係情報と、各要素の奥行き値とを原画像とともに、符号化して受信装置に伝送する。受信装置は、符号化されて伝送されてきた2次元映像信号、各要素間の位置関係情報、および各要素の奥行き値を復号化し、それらの情報に基づいて、図6(e)、(f)に示すように、視差映像としての右映像225、左映像226を生成する。

【0032】次に、視差映像を変換する際に、生じる画素の無い領域の補間方法を図7に示す。つまり、立体効果を出すために被写体像を回転あるいは移動させて立体像を作成した場合、必ずしも元の2次元表示面積と、立体像のための面積が一致するとは限らない。このため、背景画像に対して画素の無い無画部が生じることがある。

【0033】例えば、図7(a)、(b)に示すように、送信装置にて原画像231に対して、背景の映像232、家の映像233に分解された各要素を受信装置が受信し、その各要素の画像と奥行き値に従って、視差映像を生成する。すると、図7(c)に示すように、画素の無い無画部が発生する場合がある。この場合、図7(d)に示すように、家の映像233を空間軸方向で拡大することにより補間する。また、画素の無い領域を隣接画素から補間することも可能である。

【0034】図8は、この発明の第3の実施の形態について説明するための構成図である。尚、図1と同一部分は同一符号を付しその説明を省略する。この実施の形態では、図8に示すように、受信装置301において、観察者が調整する視差調整入力装置303によりユーザが入力した視差情報を基づいて、視差映像生成装置302により立体映像を生成する。これにより、観察者好みに合わせた立体映像を提供することが可能となる。つまり、ユーザは、視差情報を調整し、立体像の奥行き感を調整することができる。

【0035】図9(a)～(c)は、視差調整の一例を

説明するための図である。受信装置301には、図9(a)に示すように、スクリーン面311に対して送信装置101の演算装置102の前処理により奥行き値312を指定する。

【0036】そして、図9(b)に示すように、ユーザが奥行きの基準面を調整する場合、ユーザが視差調整入力装置303にて奥行きの基準面の調整値を入力する。すると、受信装置301の視差映像生成装置302は、奥行き値の基準面を奥行き値312から奥行き値313に変更して視差映像を生成する。これにより、送信装置101により指定される奥行き値を観察者が入力する奥行きの基準面の調整値に従って変更し、生成される映像を全体に奥に見える映像あるいは手前に見える映像に変更することができ、観察者の好みに応じた立体映像を提供することができる。

【0037】また、ユーザが奥行き量を調整する場合、ユーザが視差調整入力装置303にて奥行き量の調整量を入力する。すると、受信装置301の視差映像生成装置302は、奥行き値312を調整量に応じて増減させて奥行き値314に変更して視差映像を生成する。これにより、送信装置により指定される奥行き値を観察者が入力する奥行き量の調整量に従って変更し、生成する映像を遠近感の大きい映像あるいは遠近感の小さい映像に変更でき、観察者の好みに応じた立体映像を提供できる。

【0038】上記第1～第3の各実施の形態で説明したように、2次元映像から立体映像を生成する処理において、立体映像送信装置は、2次元映像から立体映像を生成する上で必要な2次元映像における各画素あるいは各要素の奥行き値等の付加情報を抽出する前処理を行い、この前処理で得られた付加情報と2次元映像とを符号化した信号として立体映像受信装置へ送信する。立体映像受信装置は、送信された信号を受信して2次元映像と付加情報をそれぞれ復号化し、復号化された2次元映像と付加情報とに基づいて、視差情報を用いた立体映像を生成するようになっている。

【0039】これにより、多くの演算を要する奥行き値の推定などの付加情報を抽出する処理を送信装置で非リアルタイムで行い、精度の高い付加情報を得ることができ、受信装置では、映像情報を受信した際に、リアルタイムで高精度の立体映像が生成できる。

【0040】また、送信側で得られた奥行き値等の付加情報に基づいて立体映像の生成する際に、受信側で奥行き値等の付加情報を観察者が変更できるため、観察者の好みに合わせた立体映像の提供が可能となる。

【0041】

【発明の効果】以上詳述したように、この発明によれば、立体映像受信装置のコストを下げることができ、受信した映像情報をリアルタイムで処理して立体映像を生成することができる。

## 【図面の簡単な説明】

【図1】 この発明の第1の実施の形態に係わるシステム構成図。

【図2】 この発明の第1の実施の形態の概念を説明するための説明図。

【図3】 知覚する奥行きと生成する視差映像の幾何学的関係を示す図。

【図4】 この発明の第2の実施の形態に係わるシステム構成図。

【図5】 要素毎に奥行き値を付加情報とした場合の伝送フォーマットを示す図。

【図6】 第2の実施の形態の概念を説明するための図。

【図7】 画素の無い無画部の補間を説明するための図。

【図8】 第3の実施の形態に係わる立体映像システムの構成を示す図。

【図9】 画像の奥行きの調整方法を説明するための＊

## \*図。

【図10】 従来の2次元映像から立体映像を生成するシステムの構成を示す図。

【図11】 従来の2次元映像を伝送する際の構成を説明するための図。

【図12】 従来の2次元映像を伝送する際の構成を説明するための図。

## 【符号の説明】

101(201)…送信装置

102(202)…演算装置

103(203)…符号化装置

104(204)…多重化装置

105(205、301)…受信装置

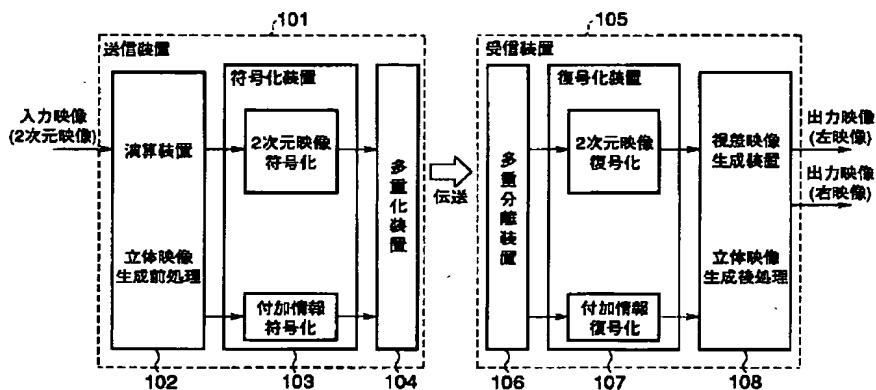
106(206)…多重分離装置

107(207)…復号化装置

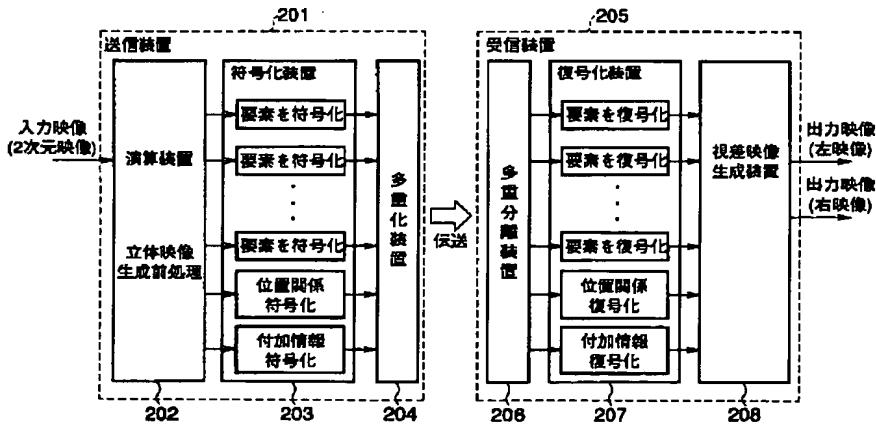
108(208、302)…視差映像生成装置

303…視差調整入力装置

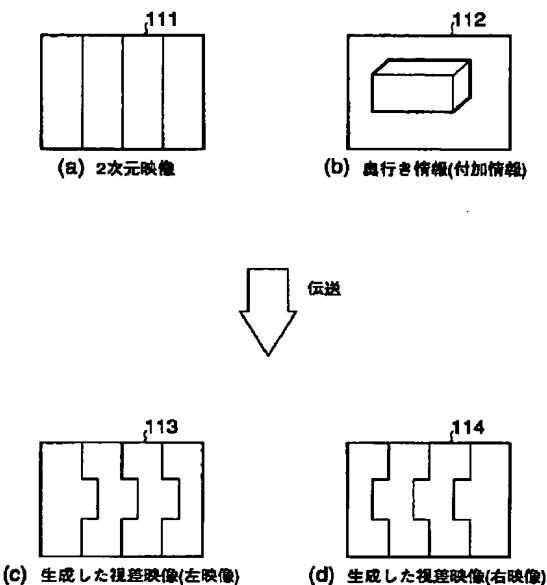
【図1】



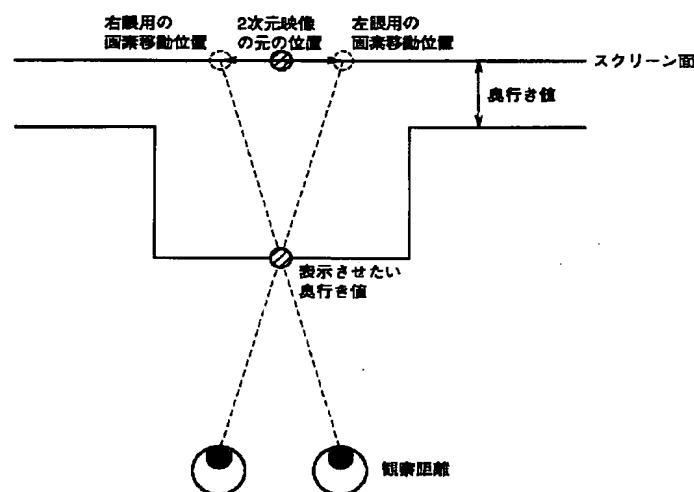
【図4】



〔図2〕



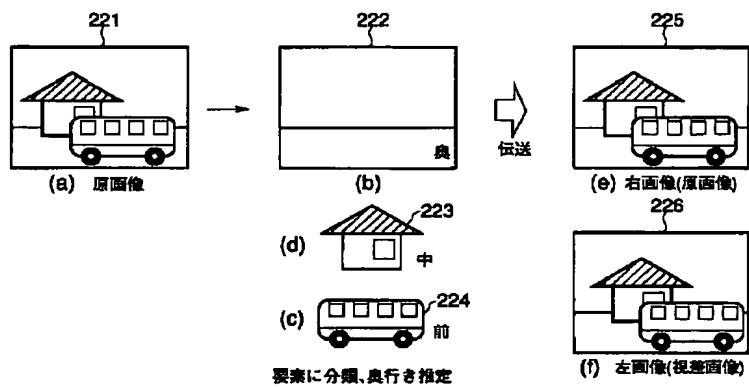
(図3)



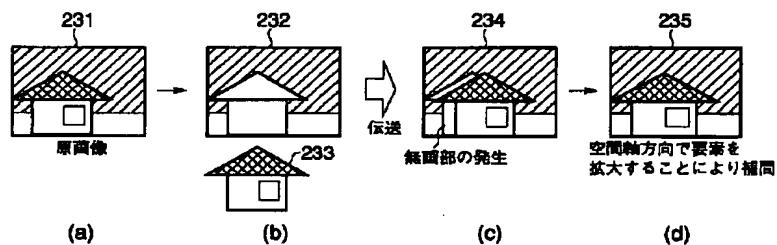
【図5】

211								
Index	name	depth	name	depth	· · · · ·		name	depth

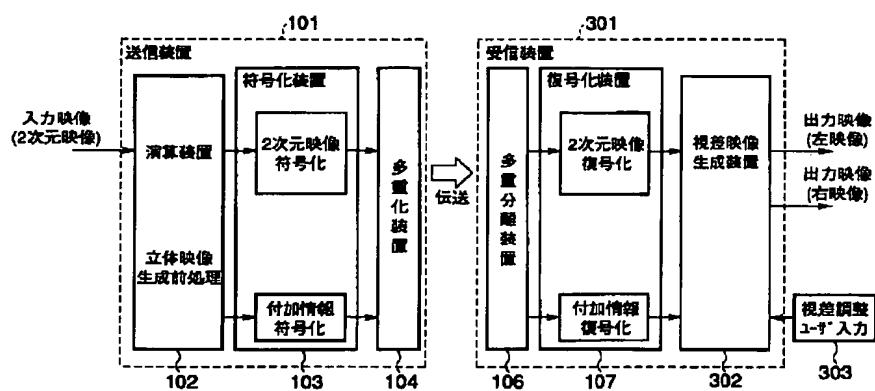
【図6】



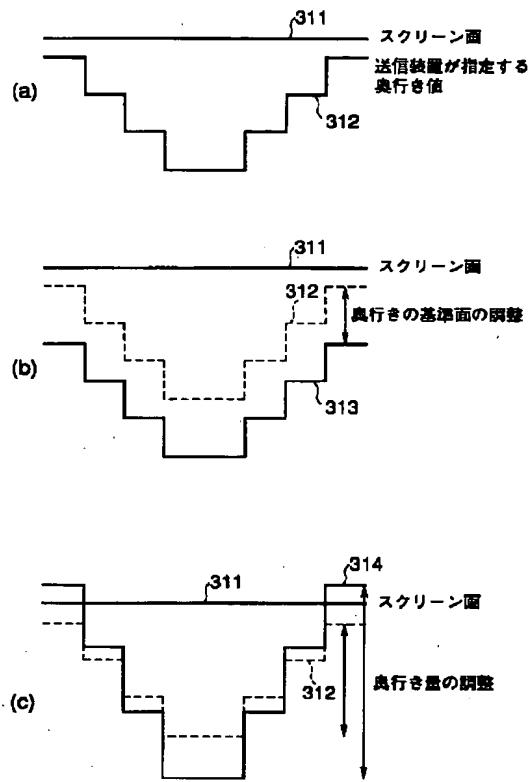
【図7】



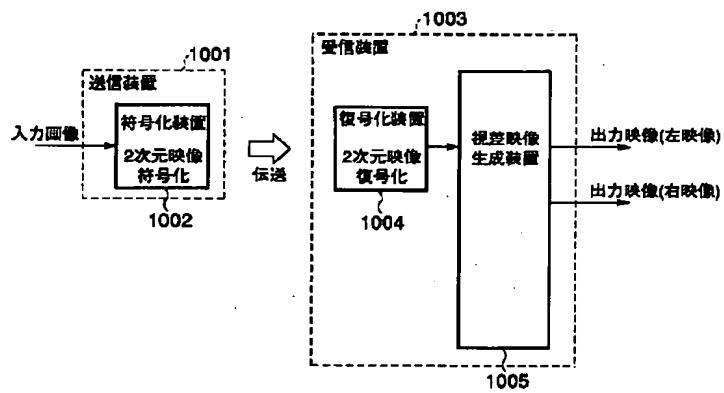
【図8】



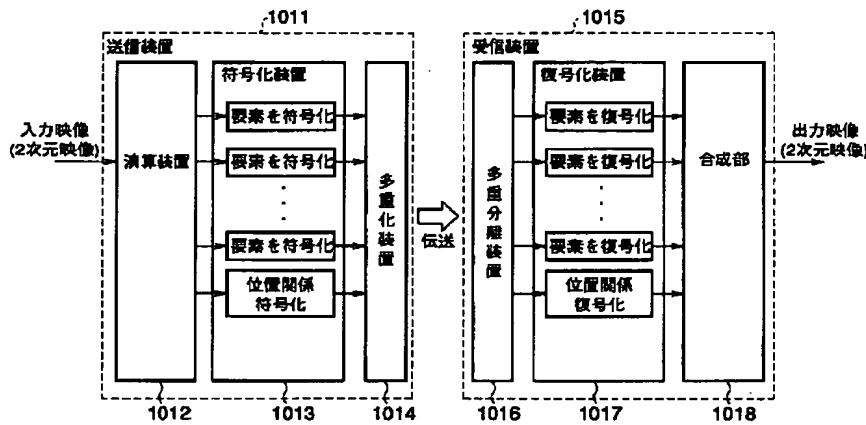
【図9】



【図10】



【図11】



【図12】

